American Journal of Epidemiology

Vol. 162, No. 11
DOI: 10.1093/aje/kwi322
Advance Access publication October 19, 2005

Practice of Epidemiology

The Shanghai Women's Health Study: Rationale, Study Design, and Baseline Characteristics

Wei Zheng[1,2], Wong-Ho Chow[3], Gong Yang[1,2], Fan Jin[4], Nathaniel Rothman[3], Aaron Blair[3], Hong-Lan Li[4], Wanqing Wen[1,2], Bu-Tian Ji[3], Qi Li[4], Xiao-Ou Shu[1,2], and Yu-Tang Gao[4]

[1] Center for Health Services Research, Department of Medicine, Vanderbilt University, Nashville, TN.
[2] Vanderbilt-Ingram Cancer Center, Vanderbilt University, Nashville, TN.
[3] Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD.
[4] Department of Epidemiology, Shanghai Cancer Institute, Shanghai, People's Republic of China.

Received for publication January 27, 2005; accepted for publication June 24, 2005.

Although cancer is a major cause of morbidity and mortality in most nations, the spectrum of cancer occurrence varies substantially worldwide. Most previous epidemiologic studies investigating cancer etiology were conducted in North American and western European countries that are relatively homogenous in terms of cancer spectrums and many lifestyle exposures. These limitations may have hindered these studies from evaluating some important etiologic hypotheses. From 1996 to 2000, the Shanghai Women's Health Study recruited 74,942 adult Chinese women from selected urban communities, with a 92% response rate. All participants completed a detailed baseline survey and anthropometrics. Approximately 88% of cohort members donated a urine sample ($n = 65{,}755$) and a blood ($n = 56{,}832$) or exfoliated buccal cell ($n = 8{,}934$) sample. Noteworthy characteristics of this cohort include low consumption of alcohol (1.9%) and use of tobacco (2.4%); high intake of fish (mean, 50.8 g/day), soy foods (mean, 142.3 g/day), and certain vegetables; low prevalence of obesity (5.1%); and nearly 100% employment outside the home. Currently, this cohort of women is being followed via biennial in-person recontact and periodic linkage to cancer and vital statistics registries. The resources from the cohort will be valuable in future studies of environmental exposures and biomarkers for the risk of cancer and other chronic diseases.

cohort studies; diet; neoplasms; nutrition; occupations

Abbreviation: SWHS, Shanghai Women's Health Study.

Although cancer is the major cause of morbidity and mortality among women in most nations, the spectrum of cancer incidence varies substantially worldwide (1). Breast, lung, and colon cancers are the three most common malignancies among Caucasian American women, accounting for over half of all cancer diagnoses in 1988–1992 (table 1) (1). These cancers, however, are much less common in many developing countries, including China (1, 2). In Shanghai, the largest city on the east coast of China, the 1988–1992 age-adjusted incidence rates per 100,000 women for breast, lung, and colon cancers were 26.5, 18.2, and 10.8, respectively, roughly 25–50 percent of the rates observed among Caucasian women in the United States (1, 2). The incidence rates of several other major cancers, including those of the corpus uteri, ovaries, bladder, and pancreas, are also substantially higher in American women than in their Chinese counterparts. On the other hand, women in China have a markedly higher incidence of cancers of the stomach, liver, and esophagus than women in the United States. Chinese immigrants to the United States have experienced a rapid change in cancer spectrum, with the incidence rates of common cancers gradually approaching rates observed for

Reprint requests to Dr. Wei Zheng, Center for Health Services Research, Vanderbilt University, Medical Center East, #6000, Nashville, TN 37232-8300 (e-mail: wei.zheng@vanderbilt.edu).

TABLE 1. Age-adjusted incidence rates (per 100,000 person-years)* of major cancers among women in urban Shanghai, China, and the United States

Cancer site (*International Classification of Diseases*, Ninth Revision code(s))	Shanghai urban area			United States (1988–1992), by race†			
	1972–1974	1988–1992	% change	Chinese	African American	White	Latino
Esophagus (150)	11.3	4.8	−57.5	0.8	3.8	1.5	0.5
Stomach (151)	23.9	21.0	−12.1	6.9	7.6	3.0	4.8
Colon (153)	5.7	10.8	89.5	21.1	27.9	20.7	15.5
Rectum (154)	6.7	7.3	9.0	10.0	8.5	9.0	7.1
Liver (155)	12.0	9.8	−18.3	6.0	2.1	1.2	2.5
Pancreas (157)	3.1	4.1	32.3	4.9	11.0	6.2	5.3
Lung (162)	18.0	18.2	1.1	24.0	44.3	40.4	22.9
Breast (174–175)	18.3	26.5	44.8	55.2	83.7	103.3	70.8
Cervix uteri (180)	26.7	3.3	−87.6	6.0	10.5	6.2	12.9
Corpus uteri (182)	2.5	3.7	48.0	11.3	11.4	18.8	13.5
Ovary (183)	4.8	5.8	20.8	7.8	9.5	13.2	9.9
Bladder (188)	2.0	1.8	−11.3	3.0	4.3	6.5	2.0
All sites (140–208)	175.3	154.3	−12.0	204.3	287.1	306.3	231.7
All sites except nonmelanoma skin cancer (173)	173.6	153.2	−11.8	204.2	286.0	305.4	231.1

* Directly age adjusted by using the world standard.
† Data from the San Francisco Bay Area of California (1).

Caucasian women (3), suggesting the importance of changing lifestyle and other environmental exposures in the etiology of these common malignancies.

Compared with women living in the United States and other Western countries, Chinese women differ substantially in many lifestyle factors. For example, traditional plant-based foods constitute a large part of the Chinese diet, including soy products, garlic, and cruciferous vegetables (4). In addition, a high percentage of Chinese people drink tea regularly. Numerous in vitro and animal studies have found that these dietary factors inhibit carcinogenesis and tumor formation (5–9). In Western cultures, these foods are consumed infrequently or at very low levels, which hinders an informative analytic evaluation. The majority of women in China have worked outside the home, facilitating investigations of occupational exposures in relation to disease risk. Cancers of the stomach, liver, and esophagus are uncommon in most Western societies (1). These cancers, however, are still among some of the most commonly diagnosed malignancies in Chinese women, although the rates have declined gradually over the years (2). It is possible that some lifestyle factors common in China may have contributed to the high risk of these cancers among Chinese women. Therefore, studies in China, with its unique patterns of lifestyles, environments, and cancer spectrum, may provide excellent opportunities to address some important etiologic hypotheses for cancer and other chronic diseases that cannot be adequately investigated in studies conducted in Western countries.

In 1996, we launched a large cohort study, the Shanghai Women's Health Study (SWHS), with an initial goal to recruit 75,000 women and collect blood and urine samples for 20,000 cohort members. Additional support was obtained from the National Cancer Institute Intramural Program to strengthen the occupational and environmental component of the study and expand biologic sample collection to all eligible cohort members. In this paper, we describe the study methodology as well as baseline characteristics of the study cohort.

MATERIALS AND METHODS

Figure 1 depicts the major components of the study. The study was conducted in seven urban communities in Shanghai with incidence rates of major cancers as well as distributions of age, sex, educational level, and occupation similar to those in the general population of urban Shanghai. Shanghai was chosen for the study because of its large and stable population with diverse exposures and the existence of a population-based cancer registry for cohort follow-up. Furthermore, over the years, through many large epidemiologic studies, we developed and tested an array of research protocols and instruments and established an experienced team of researchers in Shanghai, which greatly facilitated implementation of the cohort study.

Subject recruitment and baseline survey

A roster of all women aged 40–70 years was obtained from the resident offices in the study communities. Each eligible woman was approached by a trained interviewer and a local community health worker. After obtaining informed consent, the interviewer then provided instruction to the subject to complete a self-administered questionnaire, the first part of the baseline survey. An appointment was made for 2–3 days later to collect the completed questionnaire, conduct an

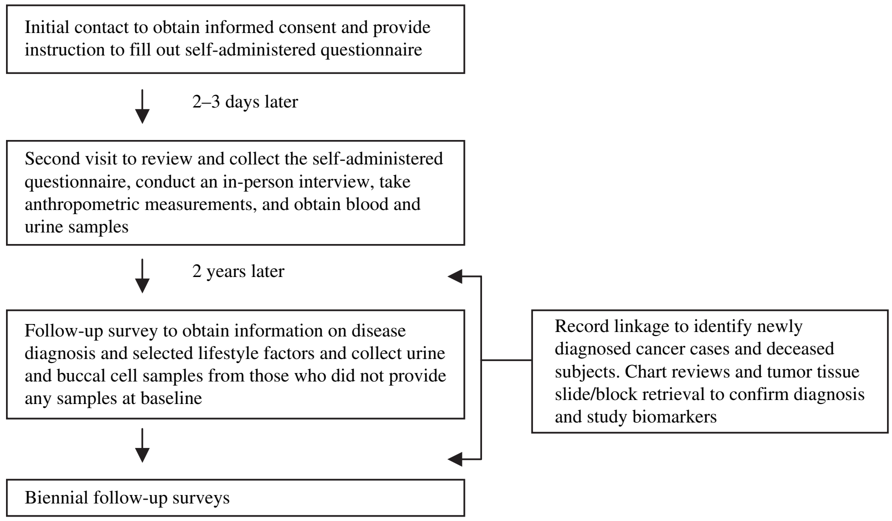

FIGURE 1. Subject recruitment, baseline survey, sample collection, and cohort follow-up for the Shanghai Women's Health Study (1996–present), Shanghai, China.

in-person interview to complete the second part of the baseline survey, and collect blood and urine samples. At the second visit, before the start of the in-person interview, the interviewer reviewed the completed self-administered questionnaire and asked a set of questions to evaluate its quality and completeness. When quality or completeness was in question, an in-person interview was conducted to complete the self-administered questionnaire. Then, the interviewer started the in-person interview, which took on average about 50 minutes. After the interview was completed, the subject's weight, standing and sitting height, and waist and hip circumferences were measured. Two measurements were taken, with a tolerance for differences of less than 1 cm for height, 0.5 cm for circumferences, and 1 kg for weight. A third measurement was taken if the difference between the first two measurements was larger than the defined tolerances.

The interviewer-administered questionnaire consists of the following five sections: 1) dietary habits, 2) reproductive history and hormone use, 3) physical activity (over the past 5 years and during the adolescent period), 4) water drinking, and 5) weight history and body measurements. In the dietary section of the questionnaire, data were obtained regarding usual dietary intake over the past 12 months and during adolescence (age 13–15 years). The questionnaire was designed to capture information on consumption of major nutrients as well as foods hypothesized to have cancer inhibitory effects, such as soy foods, allium-type vegetables, and cruciferous and dark-green, leafy vegetables. The validity and reliability of the questionnaire have been evaluated and shown to be comparable to, if not better than, those in major cohort studies conducted in other populations (10, 11). The self-administered questionnaire includes information on demographic characteristics, disease and surgery histories, personal habits (e.g., cigarette smoking, alcohol consumption, tea drinking, ginseng use, and hair dye use), menstrual history, residential history, occupational history, family history of cancer, and the subject's husband. Lifetime occupational history was obtained in the survey, including all jobs held for at least a year. For each job, the following information was obtained: name of workplace, job title, major products produced or handled, and years that each job began and ended.

Of the 81,170 eligible women who lived in the study communities during the time period of the baseline survey, 75,221 participated in the study, with a participation rate of 92.7 percent (table 2). Of those women who completed the

TABLE 2. Number of study participants and response rates for the baseline and follow-up surveys and biologic sample collection, the Shanghai Women's Health Study (1996–2000), Shanghai, China

	No. of eligible subjects	No. of participants	Response rate (%)
Baseline survey (1997–2000)	81,170	75,221*	92.7
Specimen collection			
Blood	74,942*	56,831	75.8
Urine	74,942	65,754	87.7
Exfoliated buccal cells	18,111†	8,934	49.3
Follow-up survey I (2000–2004)	74,942	74,768	99.8
Lifestyle survey	72,983‡	67,163	92.0

* Of the 75,221 study participants, 279 women who did not meet the age eligibility requirements for the study were excluded, resulting in a cohort of 74,942 subjects.

† Subjects who did not provide a blood sample at baseline and were alive at the time of the first follow-up survey when buccal cell samples were collected.

‡ Subjects who were alive at the time of the first follow-up survey.

survey, 279 were later found to be younger than age 40 or older than age 70 years at the time of the baseline interview and thus were excluded from the cohort. The remaining 74,942 women constitute the cohort. Of the 5,949 nonparticipants, 2,407 refused to participate (3.0 percent), 2,073 were absent during the study period (2.6 percent), and 1,469 (1.8 percent) were excluded for other miscellaneous reasons. The protocols used were approved by the institutional review boards of all institutions involved in this study.

Quality control procedures and data entry/editing

Interviewers were selected from among recently retired nurses and other medical professionals. Medical professionals were employed for the study largely because they have good medical knowledge and, throughout their routine work, have developed working habits to accurately attend to details. All interviewer candidates are required to complete standardized training and are certified to conduct independent surveys. A total of 92 interviewers participated in the field work at the baseline survey and sample collection, and 67 interviewers completed 98.5 percent of the interviews. All interviews were tape recorded, and 10 percent of the tapes were evaluated for interviewing quality. About 3–5 percent of subjects were recontacted by phone to evaluate the interviewers' work. All questionnaires and forms were coded twice and were double entered by different clerks. Inconsistent records were manually checked and corrected. Computer programs were also developed to check the logic and reasonable range of responses throughout the questionnaire to identify contradictory responses.

Biologic sample collection, processing, and storage

For each subject who provided consent, a 10-ml blood sample was drawn into an ethylenediaminetetraacetic acid Vacutainer tube (Becton, Dickinson and Company, Franklin Lakes, New Jersey), and a spot urine sample was collected into a sterilized 100-ml cup containing 125 mg of ascorbic acid. After collection, the samples were kept in a portable, insulated bag with ice packs (at about 0–4°C) and were processed within 6 hours for long-term storage at −70°C. A biospecimen collection form was filled out for each woman when the sample was procured. For those who did not donate a blood sample at baseline, a sample of exfoliated buccal cells was collected during the follow-up survey by using a protocol modified from the mouthwash method described by Lum and LeMarchand (12). From mouth-rinse samples, cell pellets were stored for future studies at −70°C. Of the 74,943 study participants, 56,831 (75.8 percent) donated a blood sample, and 65,754 (87.7 percent) donated a urine sample (table 2). An exfoliated buccal cell sample was collected from an additional 8,934 (49.3 percent) of the 18,111 subjects who did not provide a blood sample at baseline and were alive at the time of the follow-up survey. A computerized inventory tracking system was developed, which can quickly find the storage location of all samples for each cohort member. The biologic sample repositories for the study are equipped with appropriate alarm systems and emergency electricity backup to prevent accidental thawing.

TABLE 3. Distribution of the cohort by selected demographic factors at baseline (1996–2000), the Shanghai Women's Health Study, Shanghai, China

Baseline characteristic	No. of subjects (N = 74,942)	%
Age group (years)		
40–44	20,918	27.91
45–49	15,449	20.61
50–54	10,543	14.07
55–59	7,628	10.18
60–64	9,725	12.98
65–70	10,679	14.25
Mean age (standard deviation)	52.13 (9.08)	
Highest educational level achieved		
No formal education	8,036	10.72
Elementary school	8,147	10.87
Middle school	27,683	36.94
High school	20,890	27.87
Technical school/college or above	10,173	13.57
Unknown	13	0.02
Marital status		
Never married	658	0.88
Presently married	66,499	88.73
Widowed	5,600	7.47
Divorced/separated	813	1.08
Unknown	1,372	1.83
Ever employment*		
Professional/technical/administrative	21,389	28.54
Clerical/commercial/services	15,493	20.67
Agricultural	2,054	2.75
Manufacturing	35,367	47.37
Construction	357	0.48
Unknown	282	0.38
Menopausal status		
Premenopause	37,770	50.40
Postmenopause	37,156	49.58
Unknown	16	0.02

* A participant could have held more than one category of job during her lifetime.

Cohort follow-up and outcome ascertainment

The cohort has been followed by using a combination of record linkage and active follow-ups. Every 2 years, an interviewer visits the last known address of each living cohort member and records details of the interim health history, including cancer, cardiovascular disease, stroke, bone fracture, and other chronic diseases that have occurred since the last in-person contact. In addition to the interim health history, the survey questionnaire also includes a module to obtain information related to selected lifestyle factors. The lifestyle

TABLE 4. Selected characteristics of participants in the Shanghai Women's Health Study at baseline (1996–2000), Shanghai, China

Characteristic	%	Characteristic	$\bar{x}$ (SD)*
Prior disease history		Anthropometrics	
Malignant tumor	1.9	Body mass index†	24.0 (3.4)
Coronary heart disease	8.1	Waist-to-hip ratio	0.81 (0.05)
Stroke	2.5	Reproductive factors	
Diabetes	4.4	Age at menarche (years)	14.93 (1.74)
Lifestyle factors		No. of children‡	1.83 (1.20)
Current smoker	2.4	Age at first livebirth (years)‡	24.9 (4.1)
Current alcohol drinker	1.9	Age at natural menopause (years)§	48.1 (4.4)
Current tea drinker	29.8	Dietary nutritional factors¶	
Current use of a vitamin supplement	19.7	Energy intake (kcal/day)	1,684.5 (407.9)
Ever used hormone replacement therapy‡	2.0	Fat#	29.7 (13.5)
		Protein	67.3 (21.4)
Exercised regularly	35.5	Carbohydrate	287.0 (68.5)
Overweight (body mass index ≥25)	35.2	Red and processed meat	51.1 (36.8)
Obese (body mass index ≥30)	5.1	Poultry	15.3 (17.8)
		Fish	50.8 (45.0)
		All vegetables	298.2 (171.7)
		Cruciferous vegetables	94.8 (63.7)
		Allium-type vegetables	8.4 (10.6)
		Soy foods	142.3 (124.5)
		All fruits	270.7 (183.4)

* Values are expressed as mean (standard deviation).
† Weight (kg)/height (m)^2.
‡ Among parous women, who accounted for 96.7% of cohort members at baseline.
§ Among postmenopausal women.
¶ Unless otherwise specified, the unit is grams per day.
# Not including cooking oil.

portion of the questionnaire may vary from one survey to another. For example, a comprehensive dietary assessment was included in the first follow-up survey conducted from 2000 to 2002 and will be included again in the third follow-up survey (2007–2009), while job history was updated in the second follow-up survey (2002–2004). For subjects who have moved, the new address is requested from the Neighborhood Community Office, local police department, or the Shanghai Resident Registry.

Data routinely collected from the cancer registry and death certificates are also used to assure a timely and complete ascertainment of new cancer cases and deceased subjects in the study cohort. Cancer registration is legally mandated in Shanghai. All hospitals are required to notify the Shanghai Cancer Registry of all newly diagnosed cancer cases, and these files are linked to the cohort to identify newly diagnosed cancer cases. All possible matches are checked manually and are verified through home visits. The death certificate data from the Shanghai Vital Statistics Unit are also used to identify causes of death. For cohort members who are diagnosed with cancer, information on date and hospital of diagnosis is collected. Medical charts from the diagnostic hospital are reviewed to verify the diagnosis and collect detailed information on the pathology characteristics of the tumor. In addition, pathology slides and tumor tissue blocks are being collected for verification of cancer diagnosis and future studies of biomarkers.

The first follow-up survey was conducted from 2000 to 2002, approximately 2 years after the baseline survey. Approximately 99.8 percent of cohort members (or their next of kin, if the subjects were deceased) were interviewed, using a list of common diseases, to obtain information related to diagnoses occurring since the baseline survey. The lifestyle module of the questionnaire, including a comprehensive dietary assessment, was completed through an in-person interview for 67,163 (92.0 percent) of the 72,983 living subjects (table 2). The vast majority of cohort members (99.98 percent) continued to live in Shanghai, and most of them (79.7 percent) remained in the same neighborhood community. For nonrespondents, cancer diagnosis and vital status were identified through linkage of data from cancer and vital statistics registries.

TABLE 5. Summary of major findings from case-control studies of common cancers conducted in Shanghai, China (1996–2000)

Cancer site (reference no.(s))	Risk factors	Protective factors
Esophagus (13, 14)	Tobacco, alcohol, scalding soup/porridge, preserved vegetables, salty and deep-fried foods, irregular eating, rapid eating	Green tea, dark green–leafy/yellow vegetables, fruits, animal foods, carotene, riboflavin, vitamins C and E
Stomach (15, 16)	Preserved/salty/fried foods, scalding soup/porridge, irregular meals, rapid/binge eating, tobacco, alcohol, overweight, carbohydrates	Green tea, fruits, vegetables, soy foods, poultry, eggs, plant oil, protein, fiber, carotene, vitamins C and E
Colon (17–19)	Red meat, saturated/monosaturated fat, preserved foods, fried meat, alcohol, family history	Green tea, vegetables (cruciferous, dark green, carrots, radishes, legumes, garlic), fruits, preserved foods, dietary fiber, calcium, riboflavin, carotene, vitamins C and E, physical activity
Pancreas (18, 20, 21)	Tobacco, preserved vegetables, deep-fried/grilled/cured/smoked foods, overweight, parity, occupational exposures	Green tea, fruits, vegetables, legumes, fiber, retinol, carotene, vitamins C and E
Head and neck (22–24)	Tobacco, alcohol, salt-preserved foods, occupational exposures	Dark green/yellow vegetables, garlic, citrus fruits
Lung (25–28)	Tobacco, environmental tobacco smoke, tuberculosis and other chronic pulmonary diseases, exposure to cooking oil vapors, short menstrual cycle, occupational exposures, indoor air pollution	Fruits, vegetables, green tea
Breast (29–42)	Reproductive factors, benign breast diseases, animal foods, well-done meats, overweight, central obesity, insulin-like growth factors, C-peptide, estrogens, occupational exposures, genetic factors	Soy foods, folate, vegetable oil, vitamin E, selected fruits and vegetables (dark yellow, allium, white radishes), physical activity
Corpus uteri (43–47)	Fat and protein of animal origin, overweight, central obesity, reproductive factors	Parity, physical activity, soy food intake
Ovary (48)	Fat of animal origin, reproductive factors	Occupational physical activity, parity

RESULTS

Presented in table 3 are demographic characteristics of study participants at the baseline survey. The mean age of cohort members was 52.1 years. Nearly 80 percent had attended at least middle school. More than 88 percent of the women were married. Virtually all had ever worked outside the home, and nearly 70 percent had held at least one job with an average total duration of 12.2 years in manufacturing, agriculture, or construction. Postmenopausal women constituted approximately half of the cohort at baseline.

At baseline, nearly 2 percent of cohort members (n = 1,390) reported ever having been diagnosed with a malignant tumor in their lifetime (table 4). Among these women, breast cancer was the most common malignancy, accounting for approximately 40 percent of all prevalent cancer cases. Other common chronic diseases prevalent at baseline included coronary heart disease (8.1 percent), stroke (2.5 percent), and diabetes (4.4 percent).

Very few women in this cohort reported smoking cigarettes (2.4 percent), drinking alcohol, (1.9 percent), or using hormone replacement therapy after menopause (2.0 percent) (table 4). A high percentage of cohort members regularly drank tea (29.8 percent), took vitamin supplements (19.7 percent), or participated in leisure-time physical exercise (35.5 percent). Approximately one third of study participants had a body mass index (weight (kg)/height (m)^2) of 25 or greater, and 5.1 percent of them were obese (body mass index $\geq$30). Mean body mass index and waist-to-hip ratio were 24.0 and 0.81, respectively. Approximately 97 percent of cohort members reported having had one or more livebirths. Mean age was 24.9 years at the first livebirth, 14.9 years at menarche, and 48.1 years at menopause. Table 4 also presents mean intake levels of total energy, macronutrients, and major food groups. Noteworthy patterns include low intake of fat and high intake of fish, soy foods, and cruciferous vegetables.

DISCUSSION

Over the past two decades, we have conducted over 20 case-control and record linkage epidemiologic studies in Shanghai to investigate risk factors for cancer. The major findings for common cancers are summarized in table 5 (13–48). As in studies conducted elsewhere, tobacco smoking has been identified as the major risk factor for cancers of the lung, esophagus, kidney, oral cavity, nasal cavity, and larynx. Overweight, central obesity, physical inactivity, and reproductive history were found to be associated with the risk of several cancers, including those of the colon, breast, ovary, and corpus uteri. Certain occupations and industrial exposures, such as benzene, other solvents, asbestos, and various dusts, have been implicated in the risk of childhood cancer (49), adult leukemia (50), and cancers of the lung (28), pancreas (51), nasopharynx (52), and larynx (22). Perhaps one of the most significant contributions of the Shanghai studies thus far has been the identification of many dietary factors in relation to the risk of cancer. With the exception of esophageal cancer, intake of animal foods was, in general, associated with an increased risk of cancer, whereas plant foods were associated with a reduced risk. In particular, consumption of green tea, allium-type vegetables, and soy foods was found to be associated with reduced risks of several cancers, including those of the breast, corpus uteri, colon, esophagus, stomach, pancreas, lung, larynx, and oral cavity. These findings are promising and could be investigated further in the SWHS, which, with a cohort research design, is likely to provide more credible data than previous case-control studies.

In addition to verifying the results from previous case-control studies, the SWHS, with its extensive collection of data, provides excellent opportunities to test many new etiologic hypotheses for cancer and other chronic diseases. In particular, biologic samples were collected from the majority of SWHS cohort members, substantially enhancing our ability to assess both exogenous and endogenous exposures and to evaluate genetic factors. Biomarkers are particularly useful when the exposure cannot be assessed easily or accurately by using a survey questionnaire, as in the case of dietary intake of phytoestrogens, isothiocyanates, and folate as well as environmental exposure to polycyclic aromatic hydrocarbons. Many exposure biomarkers are aggregate measures of the level of exposure, absorption, and metabolism, thus reflecting a cumulative internalized biologic dose from multiple sources of exposure; in etiologic studies, this is a more relevant exposure assessment than that assessed by using questionnaire data. Therefore, use of biomarkers in epidemiologic studies provides added insight into the etiology of cancer.

Cohort studies have traditionally relied on a self-administered mail survey to obtain baseline exposure information (53, 54), and the quality of data may differ considerably by educational level, thus resulting in potential bias. Mail surveys may have also contributed to the low response rates commonly seen in cohort studies. When exposure patterns are related to subjects' ability and willingness to participate in studies, investigations with a low response rate may reflect only part of the exposure spectrum, thus affecting the validity of study results. Education and socioeconomic status are major determinants of study participation, particularly in self-administered surveys (55, 56). It has been documented that these two factors are closely related to many lifestyle and environmental exposures such as usual dietary and physical activity patterns (57, 58). To overcome these problems, the SWHS has made every effort to increase participation rates in both the baseline survey and cohort follow-ups. In-person interviews are used in the survey to enhance the quality of data from subjects with various educational levels. Furthermore, through repeat assessment of major lifestyle factors, such as usual dietary intake, the SWHS will provide a stable estimate of certain exposures to improve the validity and reliability of exposure data, and to provide opportunities to evaluate the timing of exposure in relation to disease risk.

In summary, over the past few years, we have successfully established a population-based cohort of 74,942 women for long-term epidemiologic studies of cancer and other chronic diseases. Currently, the cohort is being followed with a well-established and tested protocol, including record linkage with multiple sources of routinely collected data in Shanghai and biennial home visits. This study has already generated data to evaluate important etiologic hypotheses for several common diseases (59–67). Through a longer follow-up, the SWHS will provide a valuable opportunity to evaluate many important etiologic hypotheses for cancer and other chronic diseases that cannot be adequately investigated in studies conducted in Western countries.

ACKNOWLEDGMENTS

This research was supported by National Institutes of Health research grant R01 CA70867 and by Intramural Research Program contract N02 CP1101066.

The authors thank the research staff of the Shanghai Women's Health Study.

Conflict of interest: none declared.

REFERENCES

1. Parkin DM, Whelan SL, Ferlay J, et al, eds. Cancer incidence in five continents. Vol VII. Lyon, France: International Agency for Research on Cancer, 1997. (IARC scientific publication 143).
2. Jin F, Devesa SS, Chow WH, et al. Cancer incidence trends in urban Shanghai, 1972–1994: an update. Int J Cancer 1999;83: 435–40.
3. King H, Li JY, Locke FB, et al. Patterns of site-specific displacement in cancer mortality among migrants: the Chinese in the United States. Am J Public Health 1985;75:237–42.
4. Chen CM. Nutrition status of the Chinese people. Biomed Environ Sci 1996;9:81–92.
5. Lambert JD, Yang CS. Cancer chemopreventive activity and bioavailability of tea and tea polyphenols. Mutat Res 2003; 523–524:201–8.
6. Sarkar FH, Li Y. Mechanisms of cancer chemoprevention by soy isoflavone genistein. Cancer Metastasis Rev 2002;21: 265–80.
7. Conaway CC, Yang YM, Chung FL. Isothiocyanates as cancer chemopreventive agents: their biological activities and metabolism in rodents and humans. Curr Drug Metab 2002; 3:233–55.
8. Wargovich MJ, Uda N. Allium vegetables and the potential for chemoprevention of cancer. Adv Exp Med Biol 1996;401: 171–7.
9. Murillo G, Mehta RG. Cruciferous vegetables and cancer prevention. Nutr Cancer 2001;41:17–28.
10. Shu XO, Yang G, Jin F, et al. Validity and reproducibility of the food frequency questionnaire used in the Shanghai Women's Health Study. Eur J Clin Nutr 2004;58:17–23.
11. Matthews CE, Shu XO, Yang G, et al. Reproducibility and validity of the Shanghai Women's Health Study physical activity questionnaire. Am J Epidemiol 2003;158:1114–22.
12. Lum A, Le Marchand L. A simple mouthwash method for obtaining genomic DNA in molecular epidemiological studies. Cancer Epidemiol Biomarkers Prev 1998;7:719–24.
13. Gao YT, McLaughlin JK, Gridley G, et al. Risk factors for esophageal cancer in Shanghai, China. II. Role of diet and nutrients. Int J Cancer 1994;58:197–202.

14. Gao YT, McLaughlin JK, Blot WJ, et al. Reduced risk of esophageal cancer with green tea consumption. J Natl Cancer Inst 1994;86:855–8.
15. Ji BT, Chow WH, Yang G, et al. Dietary habits and stomach cancer in Shanghai, China. Int J Cancer 1998;76:659–64.
16. Ji BT, Chow WH, Yang G, et al. The influence of cigarette smoking, alcohol, and green tea consumption on the risk of carcinoma of the cardia and distal stomach in Shanghai, China. Cancer 1996;77:2449–57.
17. Yang G, Gao YT, Ji BT, et al. Dietary factors and cancer of the colon and rectum in a population-based case-control study in Shanghai. Chin J Epidemiol 1994;15:299–303.
18. Ji BT, Chow WH, Hsing AW, et al. Green tea consumption and the risk of pancreatic and colorectal cancers. Int J Cancer 1997;70:255–8.
19. Chiu BC, Ji BT, Dai Q, et al. Dietary factors and risk of colon cancer in Shanghai, China. Cancer Epidemiol Biomarkers Prev 2003;12:201–8.
20. Ji BT, Chow WH, Gridley G, et al. Dietary factors and the risk of pancreatic cancer: a case-control study in Shanghai, China. Cancer Epidemiol Biomarkers Prev 1995;4:885–93.
21. Ji BT, Hatch MC, Chow WH, et al. Anthropometric and reproductive factors and the risk of pancreatic cancer: a case-control study in Shanghai, China. Int J Cancer 1996;66:432–7.
22. Zheng W, Blot WJ, Shu XO, et al. Diet and other risk factors for laryngeal cancer in Shanghai, China. Am J Epidemiol 1992;136:178–91.
23. Zheng W, Blot WJ, Shu XO, et al. A population-based case-control study of cancers of the nasal cavity and paranasal sinuses in Shanghai. Int J Cancer 1992;52:557–61.
24. Zheng W, Blot WJ, Shu XO, et al. Risk factors for oral and pharyngeal cancer in Shanghai, with emphasis on diet. Cancer Epidemiol Biomarkers Prev 1992;1:441–8.
25. Gao YT, Blot WJ, Zheng W, et al. Lung cancer among Chinese women. Int J Cancer 1987;40:604–9.
26. Gao YT, Zheng W, Gao RN, et al. Tobacco smoking and its effect on health in China. In: O'Neill JK, Chen J, Bartsch H, eds. Relevance to human cancer of *N*-nitroso compounds, tobacco smoke and mycotoxin. Lyon, France: International Agency for Research on Cancer, 1991. (IARC scientific publication 105).
27. Zheng W, Blot WJ, Liao ML, et al. Lung cancer and prior tuberculosis infection in Shanghai. Br J Cancer 1987;56:501–4.
28. Levin LI, Zheng W, Blot WJ, et al. Occupation and lung cancer in Shanghai: a case-control study. Br J Ind Med 1988;45:450–8.
29. Gao YT, Shu XO, Dai Q, et al. Association of menstrual and reproductive factors with breast cancer risk: results from the Shanghai Breast Cancer Study. Int J Cancer 2000;87:295–300.
30. Zheng W, Jin F, Dunning LA, et al. Epidemiological study of urinary 6beta-hydroxycortisol to cortisol ratios and breast cancer risk. Cancer Epidemiol Biomarkers Prev 2001;10:237–42.
31. Shrubsole MJ, Jin F, Dai Q, et al. Dietary folate intake and breast cancer risk: results from the Shanghai Breast Cancer Study. Cancer Res 2001;61:7136–41.
32. Shu XO, Jin F, Dai Q, et al. Soyfood intake during adolescence and subsequent risk of breast cancer. Cancer Epidemiol Biomarkers Prev 2001;10:483–8.
33. Matthews CE, Shu XO, Jin F, et al. Lifetime physical activity and breast cancer risk in the Shanghai Breast Cancer Study. Br J Cancer 2001;84:994–1001.
34. Dai Q, Shu XO, Jin F, et al. Consumption of animal foods, cooking methods, and risk of breast cancer. Cancer Epidemiol Biomarkers Prev 2002;11:801–8.
35. Dai Q, Shu XO, Jin F, et al. Population-based case-control study of soyfood intake and breast cancer risk in Shanghai. Br J Cancer 2001;85:372–8.
36. Yang G, Lu G, Jin F, et al. Population-based, case-control study of blood C-peptide level and breast cancer risk. Cancer Epidemiol Biomarkers Prev 2001;10:1207–11.
37. Yu H, Jin F, Shu XO, et al. Insulin-like growth factors and breast cancer risk in Chinese women. Cancer Epidemiol Biomarkers Prev 2002;11:705–12.
38. Malin AS, Dai Q, Shu XO, et al. Intake of fruits, vegetables and selected micronutrients in relation to the risk of breast cancer. Int J Cancer 2003;105:413–18.
39. Yu H, Shu XO, Li BDL, et al. Joint effects of insulin-like growth factors and sex steroids on breast cancer risk. Cancer Epidemiol Biomarkers Prev 2003;12:1067–73.
40. Dai Q, Franke AA, Fin J, et al. Urinary excretion of phytoestrogens and risk of breast cancer among Chinese women in Shanghai. Cancer Epidemiol Biomarkers Prev 2002;11:815–21.
41. Zheng W, Gao YT, Shu XO, et al. Population-based case-control study of *CYP11A* gene polymorphism and breast cancer risk. Cancer Epidemiol Biomarkers Prev 2004;13:709–14.
42. Ren Z, Gao YT, Cai Q, et al. Genetic polymorphisms in the *IGFBP3* gene: association with breast cancer risk and blood IGFBP-3 protein levels among Chinese women. Cancer Epidemiol Biomarkers Prev 2004;13:1290–5.
43. Shu XO, Zheng W, Potischman N, et al. A population-based case-control study of dietary factors and endometrial cancer in Shanghai, People's Republic of China. Am J Epidemiol 1993;137:155–65.
44. Xu WH, Zheng W, Xiang YB, et al. Soya food intake and risk of endometrial cancer among Chinese women in Shanghai: population based case-control study. BMJ 2004;328:1285.
45. Shu XO, Hatch MC, Zheng W, et al. Physical activity and risk of endometrial cancer. Epidemiology 1993;4:342–9.
46. Xu WH, Xiang YB, Ruan ZX, et al. Menstrual and reproductive factors and endometrial cancer risk: results from a population-based case-control study in Shanghai. Int J Cancer 2004;108:613–19.
47. Shu XO, Brinton LA, Zheng W, et al. Relation of obesity and body fat distribution to endometrial cancer. Cancer Res 1992;52:3865–70.
48. Shu XO, Gao YT, Yuan JM, et al. Dietary factors and epithelial ovarian cancer. Br J Cancer 1989;59:92–6.
49. Shu XO, Gao YT, Brinton LA, et al. A population-based case-control study of childhood leukemia in Shanghai. Cancer 1988;62:635–44.
50. Adegoke OJ, Blair A, Shu XO, et al. Occupational history and exposure and the risk of adult leukemia in Shanghai. Ann Epidemiol 2003;13:485–94.
51. Ji BT, Silverman DT, Dosemeci M, et al. Occupation and pancreatic cancer risk in Shanghai, China. Am J Ind Med 1999;35:76–81.
52. Zheng W, McLaughlin JK, Gao YT, et al. Occupational risks for nasopharyngeal cancer in Shanghai. J Occup Med 1992;34:1004–7.
53. Willette W, ed. Nutritional epidemiology. Oxford, United Kingdom: Oxford University Press, 1990.
54. Cho E, Smith-Warner SA, Spiegelman D, et al. Dairy foods, calcium, and colorectal cancer: a pooled analysis of 10 cohort studies. J Natl Cancer Inst 2004;96:1015–22.
55. Turrell G, Patterson C, Oldenburg B, et al. The socioeconomic patterning of survey participation and non-response error in a multilevel study of food purchasing behaviour: area- and individual-level characteristics. Public Health Nutr 2003;6:181–9.

56. Goldberg M, Chastang JF, Leclerc A, et al. Socioeconomic, demographic, occupational, and health factors associated with participation in a long-term epidemiologic survey: a prospective study of the French GAZEL cohort and its target population. Am J Epidemiol 2001;154:373–84.
57. Winkleby MA, Jatulis DE, Frank E, et al. Socioeconomic status and health: how education, income, and occupation contribute to risk factors for cardiovascular disease. Am J Public Health 1992;82:816–20.
58. Johansson L, Thelle DS, Solvoll K, et al. Healthy dietary habits in relation to social determinants and lifestyle factors. Br J Nutr 1999;81:211–20.
59. Yang G, Shu XO, Jin F, et al. Soyfood consumption and risk of glycosuria: a cross-sectional study within the Shanghai Women's Health Study. Eur J Clin Nutr 2004;58:615–20.
60. Rosenthal AD, Shu XO, Jin F, et al. Oral contraceptive use and risk of diabetes among Chinese women. Contraception 2004; 69:251–7.
61. Rosenthal AD, Jin F, Shu XO, et al. Body fat distribution and risk of diabetes among Chinese women. Int J Obes Relat Metab Disord 2004;28:594–9.
62. Zhang X, Shu XO, Gao YT, et al. Soy food consumption is associated with lower risk of coronary heart disease in Chinese women. J Nutr 2003;133:2874–8.
63. Wen W, Gao YT, Shu XO, et al. Sociodemographic, behavioral, and reproductive factors associated with weight gain in Chinese women. Int J Obes Relat Metab Disord 2003;27: 933–40.
64. Zhang X, Shu XO, Gao YT, et al. Anthropometric predictors of coronary heart disease in Chinese women. Int J Obes Relat Metab Disord 2004;28:734–40.
65. Zhang X, Shu XO, Yang G, et al. Association of passive smoking by husbands with prevalence of stroke among Chinese women nonsmokers. Am J Epidemiol 2005;161: 213–18.
66. Yang G, Shu XO, Jin F, et al. Longitudinal study of soy food intake and blood pressure among middle-aged and elderly Chinese women. Am J Clin Nutr 2005;81:1012–17.
67. Zhang XL, Shu XO, Li HL, et al. Prospective cohort study of soy food consumption and risk of bone fracture among postmenopausal women. Arch Intern Med 2005;165: 1890–5.